

# Negotiation Outcome Classification Using Language Features

Douglas P. Twitchell · Matthew L. Jensen ·  
Douglas C. Derrick · Judee K. Burgoon ·  
Jay F. Nunamaker

Published online: 27 July 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** In this paper we discuss the relationships among negotiations, integrative and distributive speech acts, and classification of negotiation outcome. Our findings present how using automated linguistic analysis can show the trajectory of negotiations towards convergence (resolution) or divergence (non-resolution) and how these trajectories accurately classify negotiation outcomes. Consequently, we present the results of our negotiation outcome classification study, in which we use a corpus of 20 transcripts of actual face-to-face negotiations to build and test two classification models. The first model uses language features and speech acts to place negotiation utterances onto an integrative and distributive scale. The second uses that scale to classify the negotiations themselves as successful or unsuccessful at the midpoint, three-quarters of the way through, and at the end of the negotiation. Classification accuracy rates were 80, 75, and 85 % respectively.

---

D. P. Twitchell

School of Information Technology, Illinois State University, Normal, IL, USA  
e-mail: dtwitch@ilstu.edu

M. L. Jensen

Information Systems Division, Price College of Business, University of Oklahoma, Norman, OK, USA  
e-mail: mjensen@ou.edu

D. C. Derrick

School of Interdisciplinary Informatics, College of Information Science and Technology,  
University of Nebraska at Omaha, Omaha, NE, USA  
e-mail: dcderrick@unomaha.edu

J. K. Burgoon · J. F. Nunamaker (✉)

Department of Management Information Systems, Center for the Management of Information,  
University of Arizona, Tucson, AZ, USA  
e-mail: jnunamaker@cmi.arizona.edu

J. K. Burgoon

e-mail: jburgoon@cmi.arizona.edu

**Keywords** Negotiation · Negotiation outcome · Language features · Machine learning · Speech acts · Classification

## 1 Introduction

Language is a necessary and integral component of person-to-person negotiations. It encapsulates much of the information, emotion, and intentions in negotiations. Language components, including syntax (grammar), semantics (meaning), and pragmatics (intentions), are used throughout negotiations and ultimately lead to negotiation outcomes. Recent work by [Sokolova and Szpakowicz \(2007\)](#) has shown that negotiation intentions are coded in language features, that automated natural language processing can be used to extract negotiation-related language features, and that those language features, combined with machine learning techniques, can be used to predict negotiation outcomes.

In this study, we use a similar technique to arrive at negotiation outcome predictions. Using general-purpose natural language processing techniques, our method extracts a set of language features that include syntactical features such as pronouns, semantic features such as affect, and pragmatic features such as expressing appreciation. Instead of using this set of syntactical, semantic, and pragmatic features to classify negotiation utterances directly, we used them to assign scalar values to the negotiation utterances by locating them on a continuum of speech acts ranging from integrative to distributive. These speech acts and their position on the scale are then used to classify the whole negotiation as successful or unsuccessful. Although the interim step of placing utterances on an integrative/distributive continuum does not necessarily improve the model accuracy, it bases the model on established negotiation theory and gives a future user of the classification system a theoretically sound reason for the classification of the negotiation outcome. In the end, using a data set of 20 actual divorce negotiation transcripts, our two-step model is able to classify 80 % of negotiation outcomes as successful or unsuccessful at the midpoint, 75 % of negotiations at the three-quarter point, and 85 % of negotiations at the conclusion.

The paper is organized as follows. Section 2 reviews speech act theory and its relation to integrative and distributive negotiation strategies. Section 2 also reviews the [Sokolova and Szpakowicz \(2005, 2007\)](#) work on classification of negotiation outcomes and contrasts it with the study we present in this paper. Section 3 describes the research model of our study including how we move from utterances to syntactical and semantic features to speech acts and finally outcomes. We also describe the sample of negotiations we used and the results of our study. Section 4 discusses the implications of our findings and suggests future work.

## 2 Related Work

### 2.1 Speech Acts and Negotiations

Speech act theory, first described by [Austin \(1962\)](#), proposes that when a person utters something, that person is also doing something. That is, when a person utters “I will

read the paper,” she is also doing something by making a commitment to actually read the paper. Depending on the context, that commitment may range from a weak indication that the paper might be read to a formal promise that she will read the paper (or face some social consequence). The dependency on context is a fundamental tenet of speech act theory. Searle (1979) expanded the theory to say that every utterance has propositional content, the truth statement that the utterance expresses, and illocutionary force, what the speaker does when speaking. For example, the utterance “will you read the paper?” has the same propositional content as the example above (the paper being read by someone), but the speakers are performing different actions with each utterance. In the former, the speaker is committing herself to reading the paper; in the latter the speaker is requesting the listener to commit herself to reading the paper. In the same work, Searle (1979) formalized a taxonomy of speech acts which included such acts as statements, apologies, and promises.

Speech acts have been used as a means for understanding and formalizing conversations and negotiations. Winograd and Flores (1986) used speech act theory as a basis for creating a system that structured user’s conversations in a business setting so that when a user promised another user to do something, that promise (the speech act) was tracked by the system. In applying speech act theory to negotiations, Chang and Woo (1994) designed an online negotiation system that restricted negotiating parties who used the system to negotiation-specific speech acts called negotiation acts such as *make claim*, *offer compromise*, and *dissent*. These parties, which need not be human, generate their own content and attach one of these speech act labels to it. This kind of structured system allows computers to negotiate with each other or with humans and have each part of the negotiation be machine-readable. Carberry and Lambert (1999) also use speech acts as a basis for modeling negotiations to the end of creating robust natural language consultation systems.

In the aggregate, individual negotiation-specific speech acts uttered by a party in a negotiation become negotiation strategies. One popular negotiation framework divides these negotiation strategies into two kinds: distributive and integrative (Putnam 1990; Sebenius 1992; Walton and McKersie 1991). Distributive strategies are those that seek attainment of goals for one party that are in conflict with the goals of the other party. Integrative strategies are those that seek attainment of goals for a party that are not in conflict with the goals of the other party. In other words, integrative strategies seek compromise and win-win outcomes while distributive strategies seek winner-take-all outcomes. Individual utterances in a negotiation can be said to have an integrative/distributive trajectory in that they lead to either integrative or distributive strategies. Donohue and Roberto (1996) developed a coding scheme to place speech act categories along a scale ranging from integrative to distributive, thus allowing quantification of qualitative acts. For example, at the most integrative end of the scale is the speech act *Comply*, which indicates agreeing with another’s position. At the other end of the scale is the speech act *Threat to take action*, in which one of the parties threatens punitive action. Although Donohue and Roberto (1996) do not call their categories speech acts, they have the hallmarks of speech act categories including a different illocutionary force for each category. The categories along with the associated integrative/distributive scale can be found in Table 2 in the methodology section below.

## 2.2 Classifying Negotiation Outcomes

With the advent of e-negotiations, there has been an associated interest in analyzing text produced by online negotiation systems. One area of interest is classifying negotiation outcomes or labeling outcomes as successful or unsuccessful. [Sokolova and Szpakowicz \(2005, 2007\)](#) published a series of two studies that use the text of negotiations from the Inspire system ([Kersten and Noronha 1999](#)) to classify the entirety of each negotiation into successful and unsuccessful categories. Basing their work on Leech and Svartvik's (2002) communicative grammar of English, they identify words and phrases that are associated with various negotiation techniques (e.g., commands, requests, advice). For example, in both studies ([Sokolova and Szpakowicz 2005, 2007](#)), the researchers identify the phrase "you should [Main Verb]" as "advice," where [Main Verb] is replaced with a verb from several possible categories (e.g., know, reply, continue). Main verb categories include activity, communication, cognition, event, perception, attitude, process, and a state of having or being and they list a number of such phrases as significant to negotiation. They then count the existence of these phrases (or components of them) as binary language features that feed into a machine-learning algorithm for classification. The first study ([Sokolova and Szpakowicz 2005](#)) analyzed the entire negotiation and attained an accuracy of 74.5 % using decision trees. The second study ([Sokolova and Szpakowicz 2007](#)) attempted to classify the outcome of the entire negotiation based on only features from the first half of the negotiation and attained an accuracy of 73.8 %.

The study we present in this paper expands on the work by [Sokolova and Szpakowicz \(2005, 2007\)](#) in three important ways. First, the language features we chose to use are general purpose features that have been used for other purposes such as detecting deception in online text ([Zhou et al. 2004a,b](#)) using Linguistic Deception Cues (LDC), profiling online conversations ([Twitchell and Nunamaker 2004](#)) using Dialog Act Modeling (DAM), and as far afield as predicting the completeness of the bereavement process after the death of a loved-one ([Pennebaker et al. 1997](#)) using Linguistic Inquiry and Word Count (LIWC) ([Pennebaker et al. 2007](#)). These features were not developed with only negotiation in mind; instead, the previously used linguistic features and the software used to extract them were readily available.

Second, instead of using the existence of phrases or words as features for the outcome classification, we introduce the intermediate step of classifying negotiation utterances as integrative or distributive speech acts and then using those speech acts and their integrative/distributive trajectories to classify negotiation outcomes. As will be shown, this technique may not necessarily improve accuracy, but it does follow established theory in negotiation analysis and allows users of this classification methodology to better understand the reasons for the outcome classification. [Sokolova and Szpakowicz \(2005, 2007\)](#) have demonstrated a direct link between syntactic and semantic features and negotiation outcome. The integrative/distributive negotiation framework discussed earlier ([Putnam 1990](#); [Sebenius 1992](#); [Walton and McKersie 1991](#)) established a link between integrative and distributive statements and negotiation outcome, providing a theoretical framework for understanding why negotiations are successful or not. The technique described in this paper builds on both approaches by using syntactic and semantic features to score utterances on an

integrative/distributive scale. The aggregate integrative/distributive trajectory can be used as a measure to visualize the progress of the negotiation, or, as we show, the trajectory can be used as a basis for classification of negotiation outcome. The intermediate step of scoring utterances before classifying the negotiation outcome adds complexity, but avoids the common problem of incomprehensible or “black box” classification (see Freitas et al. 2010 for a discussion of this problem in the biology field) and gives the technique an established theoretical grounding. We describe the technique in detail in the next section.

Finally, we use free-form, face-to-face negotiation transcripts rather than text data produced during a negotiation guided by a structured, on-line negotiation system. The transcripts were taken from actual face-to-face divorce mediation where former spouses negotiated the terms of their divorce.

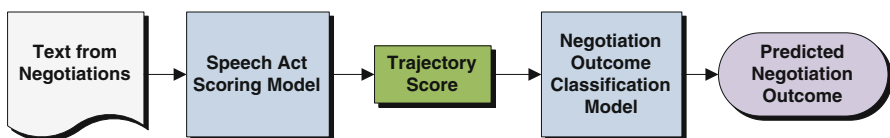
### 3 Methodology

#### 3.1 Research Model

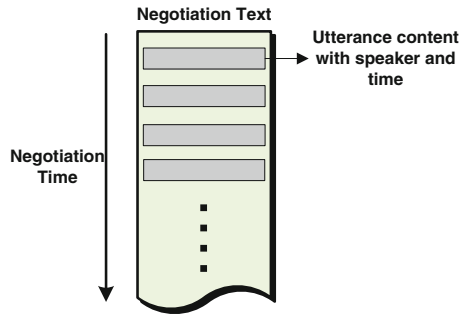
Our methodology uses the linguistic features of individual utterances to predict negotiation trajectories, which are then used to predict the outcome of a whole negotiation. We define negotiation trajectory based on integrative or distributive utterances made in the course of the interaction. If a speech act is more integrative (i.e., seeking consensus), then the trajectory is viewed as progressing towards successful resolution. If the speech act is more distributive (i.e., divisive), then the opposite would be true. In order to predict the trajectory and outcome, we use the multi-step process shown in Fig. 1 below. This process is based on automated analysis of negotiation text and transforms raw text into a current negotiation trajectory and finally a predictive outcome.

The research model begins with the raw negotiation text and requires a linguistic analysis and machine-learning-based scoring model to give each individual utterance an integrative/distributive trajectory score. A negotiation text is comprised of individual turns at talk made by the negotiating parties and often a negotiation mediator. Figure 2 shows the conceptual content of a negotiation text. The utterances are listed chronologically and include a speaker label. Time information is in the form of chronological turn number (i.e., the first speaker has turn number 1, the second, turn number 2, and so on).

The model that produces the trajectory score is created using a manually-coded training set in which each utterance is assigned a speech act category. The training set is drawn from the negotiation text and is examined by two trained coders who



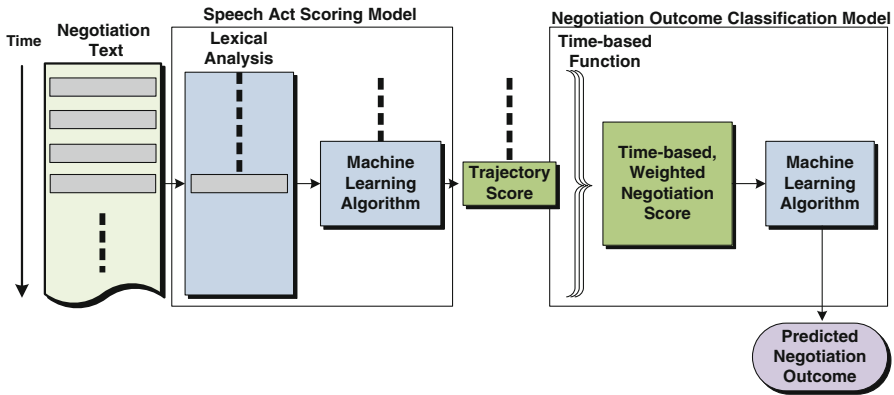
**Fig. 1** From text to trajectory and prediction



**Fig. 2** Structure of negotiation text

code the speech act categories for each utterance in the sample (see Table 2 for the speech acts and their integrative/distributive scores). Each category is mapped to an integrative/distributive score based on a one-to-one mapping and these scores are used to train the speech act scoring model. The scoring model is based on semantics (e.g., affect ratio that expresses emotion, imagery used, negative words), syntax (e.g., average word length, average sentence length), and pragmatics (e.g., general speech act categories such as statement, opinion, acknowledgement, and appreciation) and it treats all utterances discretely. Thus, the scoring model generates a single score for each utterance during the negotiation. For example, the scoring model would rate distributive utterances such as rejecting another's demand or issuing threats highly on the integrative/distributive scale (e.g., 7 or 8) and integrative utterances such as complying with another's demand lowly on the integrative/distributive scale (e.g., 1).

A second machine-learning-based classification model uses these trajectory scores to predict the overall outcome of the negotiation. We dichotomized a negotiation outcome as either a positive resolution where an agreement is reached or as a negative resolution where no accord is achieved. Both parties need not view the agreement positively for our model to view the negotiation as a success; an agreement of any form is considered successful. Game theory-based research on negotiations has shown the later a statement appears in a negotiation, the greater effect it has on the outcome (Bartos 1964). This effect is particularly true for explicit threats. Sinaceur and Neale (2005) showed that implicit threats were most effective in inducing concessions if they were early in the negotiation. Explicit threats, the kind more likely to be detected using the linguistic features used in this study, had more effect on concessions and the outcome when they were late in the negotiation. Furthermore, research on the stages of negotiations indicates that the critical stages of negotiation—those that lead to resolution—occur later in the negotiation rather than earlier (e.g., Donohue et al. 1991). Based on this research, we apply a time-based weight function after each utterance has been scored that calculates an overall trajectory score for the entire negotiation but with increasingly greater weight given to more recent utterances. Each utterance contributes to the overall negotiation score, but more recent utterances are given greater consideration than earlier ones. The resulting negotiation score is an effective predictor of the outcome



**Fig. 3** Detailed overview of process for predicting negotiation outcome

of the negotiation. Figure 3 below shows a more detailed look at the overall process.

### 3.2 Dataset

To validate our research approach, we applied the approach to 20 actual, high-stakes negotiations. The negotiations were collected during divorce mediations and involved a husband, wife, and mediator in every case. The topics of the negotiations consistently revolved around custody issues, division of resources, and child support. In these disputes, the presiding judge in the case recommended to the parties that they attempt mediation to resolve their issues. Each negotiation was recorded and transcribed. For purposes of this research, all names and places were concealed to preserve the privacy of the parties.

Of the 20 cases, eight were successfully resolved during the negotiation. The unit of analysis of this work is the utterance, meaning an uninterrupted speaking turn (Donohue and Roberto 1996). During the negotiations, the parties, including the mediator, made a total of 7,199 utterances. The distribution of utterances across negotiations and speakers is shown in Table 1.

### 3.3 Speech Act Scoring Model

To create a model for scoring individual utterances on an integrative/distributive scale, 100 consecutive statements were selected randomly from each negotiation.<sup>1</sup> Case number 11 had only 68 utterances, so all utterances from this case were included in the sample. The resultant sample included 1,968 utterances and represents 27.3% of the total utterances.

<sup>1</sup> The samples were selected by randomly selecting a statement that had 99 statements following it. The sample contained the randomly selected first statement and the consecutive 99 statements.

**Table 1** Details of the 20 negotiations studied

Case number	Resolution	Statements by husband	Statements by mediator	Statements by wife	Total
1	Yes	384	422	434	1240
2	Yes	30	54	43	127
3	No	176	151	151	478
4	Yes	94	109	79	282
5	No	133	155	149	437
6	No	63	70	77	210
7	No	52	56	59	167
8	No	72	69	43	184
9	No	93	89	104	286
10	Yes	153	155	151	459
11	No	29	30	9	68
12	Yes	79	141	159	379
13	No	51	63	64	178
14	No	128	98	111	337
15	No	91	86	63	240
16	No	193	277	166	636
17	No	212	177	134	523
18	Yes	65	94	69	228
19	Yes	57	77	90	224
20	Yes	157	212	147	516
Total	8 Resolved	2,312	2,585	2,302	7,199

Two trained coders reviewed the utterances and scored each utterance using a coding scheme developed by [Donohue and Roberto \(1996\)](#). The coding scheme allowed each coder to judge utterances and categorize them into speech act categories. These categories have a one-to-one mapping to an integrative/distributive rating. Highly integrative utterances (e.g., complying with requests) received low scores on the integrative/distributive scale. Utterances such as rejecting others' demands and threats to take action received high scores on the integrative/distributive scale. The integrative/distributive ratings were designed "along a continuum of integrative-distributive orientations" ([Donohue and Roberto \(1996, p. 216\)](#)); therefore following Donohue and Roberto, we treated the integrative/distributive score as a continuous variable for comparison, model creation, and tracking over time.

[Donohue and Roberto \(1996\)](#) note that it is possible for a single utterance to have multiple thought units (e.g., a single utterance may reject another's offer and make a demand). The vast majority of utterances in our sample did not contain multiple thought units. However, to minimize this possibility, the coders followed Donohue and Roberto's guidance to characterize an utterance as the most distributive when



**Table 2** Speech act coding scheme for individual utterances adapted from Donohue and Roberto's integrative/distributive coding scheme

Rating	Label	Description	Example
1	No code	Statements are not codable	"hmh"
	Comply	Agreeing with other's position	"Yeah that's what we agreed on"; "that would be okay"
2	Integrate	Proposing flexibility or giving support to the other's position	"You don't have to go exactly by that if you see it differently"
3	Question of fact	Asking the other about the situation, status of particular events, etc	"What happened Monday and Tuesday?"
4	Statement of fact	Stating a fact or perception about the situation	"She's sixteen going on nine she has problems and she needs help"
5	Statement of demand	Demanding the other comply with a request	"I want him to see the baby"
6	Avoidance	Attempts to move away from the conflict	"It does and it doesn't all right?"
7	Reject other's demand	Denying the other's demand	"No way no way every other weekend is sufficient"
8	Threat to take action	Promises to take punitive action	"Go to court"

multiple characterizations existed. The adapted coding scheme is shown in Table 2.<sup>2</sup> The utterances in the sample covered all categories and the proportions of utterances in each category are similar to datasets where more utterances have been coded (Donohue and Roberto 1996).

The reliability among coders on the integrative/distributive scale was lower than desired, but reasonable for exploratory research with a complex, subjective coding scheme (Kendall's  $\tau-b = 0.61$ ) (Hair et al. 2006; Landis and Koch 1977).<sup>3</sup> Coding

<sup>2</sup> Donohue and Roberto's original coding included a level for question of demand between statement of demand and avoidance. Question of demand was originally included in the coding scheme. However, the coders experienced difficulty separating question of fact from question of demand causing unacceptable reliability. Therefore, the only one category representing questions was retained.

<sup>3</sup> As noted by Hair et al. (2006), it is desirable that reliability among coders be at least 0.70. However, Hair et al. also note that highly subjective, exploratory scales produce lower reliability, but poor reliability should not result in discarding the data. In addition, Landis and Koch (1977) suggest that for highly subjective, exploratory research such as this reliabilities between 0.60 and 0.79 indicate *substantial agreement*.

**Table 3** Distribution of statements across speech act categories

Integrative/distributive score	Sample statements (%)
No code	5.0
1	9.1
2	0.4
3	14.6
4	51.6
5	12.8
6	0.6
7	4.4
8	1.5

inconsistencies were resolved in three steps. First, statements that were labeled as not codable by either coder were labeled with *No Code*. Second, differences of 1 or 2 between coders were resolved by adopting the more distributive rating (again, consistent with Donohue and Roberto). For example, if one coder rated an utterance with a 1 and another coder rated the same utterance as a 2, the more distributive code (the 2) was adopted. Finally, differences of 3 or greater between coders were manually inspected and resolved by the authors. Only 13.3 % of the utterances needed to be resolved by manual inspection. Most of the inconsistencies arose because of confusion between attempts to *Integrate* and *Avoidance* and between *Comply* and *Statement of Fact*. The distribution of coded utterances is shown in Table 3. Of the 1,968 utterances in the sample, 98 (5.0 %) were labeled as *No Code* and were excluded from the training data.

With the training sample manually coded, a scoring model designed to estimate the integrative/distributive score using lexical features was created. The sample utterances were automatically parsed and tagged using the LDC (Zhou et al. 2004a), the Whissell dictionary for affective words (Whissell 1989; Whissell et al. 1986), the LIWC (Pennebaker et al. 2007), and DAM (Stolcke et al. 2000).

The automatic parsing and tagging produced 148 candidate features that might be included in a model estimating the integrative/distributive score for a single utterance. As the sample of coded utterances numbered only 1,870, a feature reduction step was adopted. For feature reduction, a best first subset evaluation technique was adopted which considered each feature's individual predictive ability, but reduced the redundancy among them (Hall 1998; Hall and Smith 1998). Using this feature reduction technique, a total of 14 features were selected for inclusion in the model. These features are described in Table 4.

A nonlinear relationship was anticipated between the features selected for estimating the integrative/distributive score. For example, elevated affect could be present in a *threat to take action* (coded as 8) or in an utterance of *demand* (coded as 5) but totally absent from an utterance of *avoidance* (coded as 6). Furthermore, not all of the features are independent of one another. Therefore, support vector regression (Shevade et al. 2000; Smola and Schoelkopf 2004) was selected as the method for model creation as it is not constrained by assumptions of linearity or independence (Witten and Frank

**Table 4** Linguistic features used to score utterances

Feature	Source <sup>†</sup>	Description
I*	LIWC	First person singular pronouns such as <i>I, me</i>
Future*	LIWC	Verbs in future tense such as <i>will</i>
Job*	LIWC	Words concerning employment
Lexical diversity*	LDC	Unique words
Imagery*	LDC/Whissell	The amount of imagery expressed by words in the affect dictionary
Unpleasant words*	LDC/Whissell	Words in the affect dictionary whose unpleasantness rating is greater than a 1 standard deviation from the mean
Tag question <sup>‡</sup>	DAM	Questions added to the end of a sentence seeking agreement such as ..., <i>you know?</i>
Agree/accept <sup>‡</sup>	DAM	An agreement or acceptance of another's statement or opinion
Acknowledgement <sup>‡</sup>	DAM	Continuers such as <i>uh-huh</i>
Appreciation <sup>‡</sup>	DAM	Expressing appreciation (e.g., <i>thank you</i> )
Down playing <sup>‡</sup>	DAM	Mild, friendly disagreement in response to apologies or compliments (e.g., <i>that's ok</i> )
Acknowledge-answer <sup>‡</sup>	DAM	Acknowledgment of an answer (e.g., <i>Oh, okay</i> or <i>I see</i> )
Yes answer <sup>‡</sup>	DAM	Affirmative answers such as <i>yes</i>
Statement of opinion <sup>‡</sup>	DAM	Statements that include opinion qualifiers such as <i>I think</i> or <i>I believe</i>

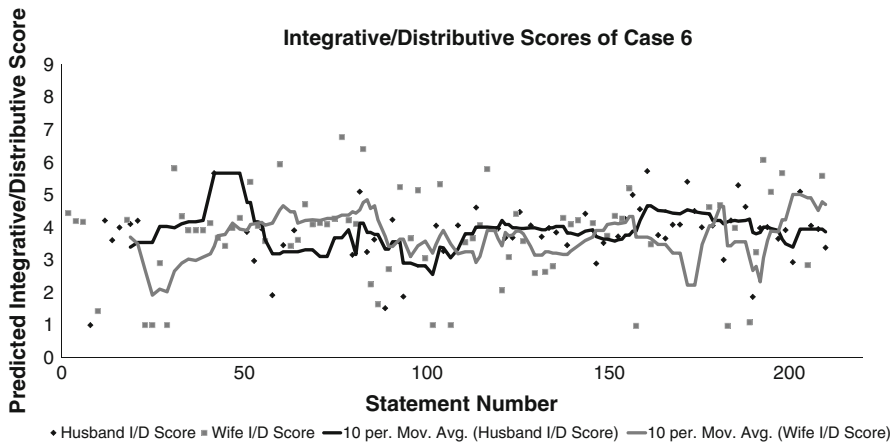
\* Normalized to the number of words in the message

<sup>†</sup> LIWC Linguistic Inquiry and Word Count (Pennebaker et al. 2007), LDC Linguistic Deception Cues (Zhou et al. 2004a,b), Whissell Whissell affect dictionary (Whissell 1989), DAM Dialog Act Modeling (Stolcke et al. 2000)

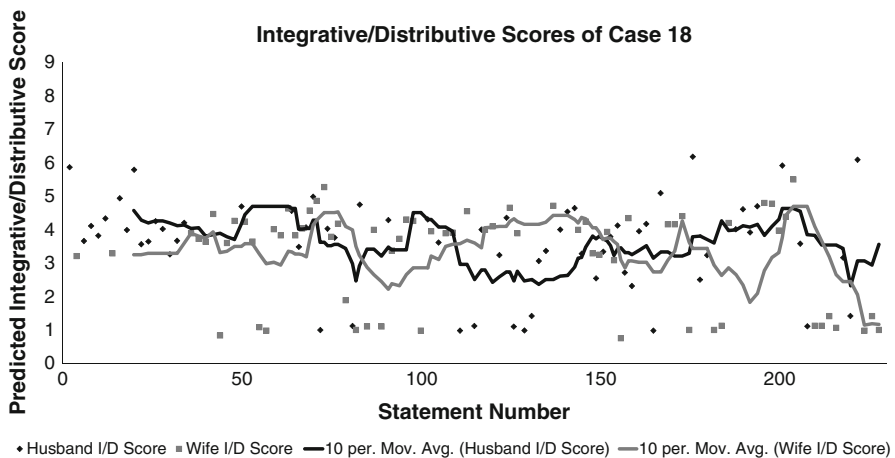
<sup>‡</sup> The DAM features are probabilities that the utterance is in the given speech act category

2005). For support vector regression, Smola and Schoelkopf's (2004) sequential minimal optimization algorithm was utilized. Using the reduced feature set (14 features), a model was created to estimate an integrative/distributive score for each utterance in the sample. The resultant model explained just under half of the variance in a cross-validated test of model performance (10-fold cross-validation;  $r=0.52$ ; Mean absolute error = 0.67; Root mean squared error = 1.18). This performance compares favorably with the ZeroR rule of assigning a score of 3.90 to all statements ( $r = -0.04$ ; Mean absolute error = 0.85; Root mean squared error = 1.36).

With the model trained on the sample of utterances, the model was then applied to all of the divorce mediation utterances. As the sample of utterances used to train the model was taken from the larger dataset, a similar correlation coefficient and mean absolute error measure were expected for the integrative/distributive scores assigned to the remaining utterances. Plots of the integrative/distributive scores from utterances in unsuccessful and successful negotiations are shown in Fig. 4 (unsuccessful) and Fig. 5 (successful). In addition to the plotted scores, a line averaging the previous 10 scores from the utterances is also shown. Only the scores from the husband and wife are displayed.



**Fig. 4** An example of a negotiation without a resolution



**Fig. 5** An example of a negotiation with a resolution

### 3.4 Negotiation Outcome Classification Model

Following the research approach, outcomes of the divorce mediations were estimated using the integrative/distributive scores from the individual utterances. This work adopts a “negotiation in stages” view, meaning that negotiations traverse multiple stages before coming to a resolution or impasse (e.g., [Donohue et al. 1991](#)). Although there is some dispute in the precise stages a negotiation must progress through (e.g., [Putnam and Jones 1990](#)), it is important to note that the more critical exchanges (offer-counteroffer, reconciliation, etc.) occur later in negotiations. Additionally, when gauging the status of a negotiation, the most recent utterances (not early utterances) provide the greatest insight to the likelihood that the parties will come to an agreement. Thus, in estimating the outcomes of negotiations, a weighting mechanism that considers latter utterances more influential than earlier utterances was used. A normalized

exponential weighting function was selected that largely discounts the opening utterances of a negotiation in favor of the recent utterances. This weighting function is shown in Eq. 1, where  $x_i$  is the utterance number at line  $i$ , and  $x_{total}$  is the total number of utterances in the interaction. The weighting function provides a weight for each utterance between 0 and 1, with latter utterances being weighed more heavily than previous utterances.

$$\frac{x_i^3}{x_{total}^3} \quad (1)$$

In estimating the outcomes of negotiations, the utterances of the mediator were excluded from the analysis. Exclusion of the mediator utterances was performed for two reasons. First, the mediator does not come to any agreement or resolution; only the parties in the negotiation can reach an agreement. The mediator simply acts as a facilitator without true power over the outcome of the negotiation. Second, the role of the mediator in this environment is to function as a bridge between the two parties. The nature of the mediator's responsibility encourages bridging and integrative utterances as the mediator seeks common ground between the two parties. This may inflate the number of integrative utterances not attributable to the parties who are actually engaged in the negotiation.

Also, the length of negotiations was originally thought to be related to negotiation outcomes. This notion was tested via a Pearson correlation to see if negotiation length (number of utterances) needed to be included as a control variable in estimating the outcome of the negotiations. With the 20 divorce mediation transcripts, negotiation length was not found to be significantly correlated with successful outcome [ $r(20) = 0.23$ ,  $p = 0.32$ ]. Therefore, negotiation length was not included in the negotiation outcome model.

To test the potential for the integrative/distributive score predicting the outcome of the negotiation, we conducted three tests. The tests examined the possibility of predicting the outcome of the negotiation using half, three-quarters and all of the utterances from each negotiation. In each test, a prediction model was created using a simple naïve Bayes classifier. The simple naïve Bayes classifier effectively handles the unequal occurrence of resolved and unresolved negotiations and its calculation is simple, yet robust (Witten and Frank 2005). The average weighted integrative/distributive score was used as the single predictor of negotiation outcome in each test.

In the first test examining the first half of the utterances, the simple naïve Bayes classifier correctly predicted the outcome of 80% of the negotiations (ROC Area=0.813; 10-fold cross-validation). This performance is better than the ZeroR rule of classifying all negotiations as unsuccessful, which would yield an accuracy rate of 60%. In the second test using three-quarters of the negotiation utterances, the classifier achieved an accuracy rate of 75% (ROC Area=0.786; 10-fold cross-validation). Using all of the utterances, the classifier was able to correctly classify 85% of all the negotiations in a cross-validated test of model performance (ROC Area=0.844; 10-fold cross-validation). The classification results are summarized in Table 5.

**Table 5** Confusion matrices resulting from 10-fold cross-validation of negotiation outcomes

Test	Actual	Predicted	
		Resolved (%)	Not resolved (%)
First half of utterances	Resolved	75.0	25.0
	Not resolved	16.7	83.3
First three-quarters of utterances	Resolved	62.5	37.5
	Not resolved	16.7	83.3
All utterances	Resolved	87.5	12.5
	Not resolved	16.7	83.3

## 4 Discussion and Conclusion

We hypothesized that the makeup of individual utterances in negotiations can be used to predict the likelihood that a negotiation will reach an agreement. This approach focuses on objectively measuring the integrative or distributive nature of each individual communication act and giving greater weight to the most recent statements. Others, such as [Sokolova and Szpakowicz \(2005, 2007\)](#), have analyzed complete negotiation texts and predicted negotiation outcome, but to our knowledge this is the first time that, using speech act theory as a guide, individual utterances have been used to provide a measure of integrative/distributive trajectory, which was then used to predict whether negotiations came to agreement. This important research distinction and the resulting findings have several important implications and limitations that will be discussed below.

There are two main findings from this study. First, utterances can be automatically scored on a scale from integrative to distributive. Second, the resulting weighted integrative/distributive scores (trajectories) for each speech act can be used to predict negotiation outcome. We tested our theory using field data of transcripts from 20 divorce negotiations. These transcripts provided a good dataset because they involve high-stakes interactions with real consequences and were not generated in a laboratory. These negotiations provided 7,199 utterances for analysis in testing our theory.

The overall accuracy of 85% using all utterances was higher than ([Sokolova and Szpakowicz 2005](#)) results of 74.5%, but since the studies used different kinds of data (free-form transcripts of spoken conversations vs. structured written text) and different amounts of data (40 vs. 5,500 participants), the results are not directly comparable. Instead, in the future, we would like to see both methods tested on the same data. Our results should aid in guiding future model development.

One of the most important implications of the research is that interactions can be objectively measured as integrative or distributive and that speech acts can be used to predict agreement or convergence. Similar to results presented by [Sokolova and Szpakowicz \(2007\)](#), classification of negotiation outcome based on the first half of the negotiation yielded high classification accuracy (relative to the Zero-R). Predicting negotiation resolution at three-quarters of the way through each negotiation also yielded a relatively high accuracy rate. This finding, based on actual negotiations,

underscores the potential for language features to be used in estimating the likelihood for a negotiation's success. The ability to measure the likelihood of negotiation success would be of particular interest in on-going and real-time negotiations. Using this process, negotiation parties would be able to determine the integrative/distributive nature of each speech act and then predict the likelihood of agreement. Such a measurement would be advantageous in negotiations. For example, if the model starts to show divergence, it would allow the parties to take corrective actions in order to help move back towards agreement. It would also allow them to objectively measure and monitor their own tone and utterances as well as the other party's. The trajectory of each utterance can show the current state of integration or distribution of the interaction as well as the overall nature of the interaction to date. Since the process is automated, it is possible to use a computer support system to aid in, monitor and, perhaps as Figures 4 and 5 illustrate, visualize the negotiation process.

#### 4.1 Limitations

As mentioned previously, our model only predicts if an agreement will be reached. We dichotomized the outcome as either a positive resolution where an agreement is reached or as a negative resolution where no accord is achieved. Although each party may view the results of the negotiation differently (e.g., more positively or negatively), we dichotomized the results because if agreement is achieved then the terms are at least acceptable to each participant. Therefore, there may be a discrepancy between how the participants view the outcome of the negotiation and our process. We do not account for the "quality" of the agreement only that the agreement was reached. For example, some of the parties may have agreed in the negotiation, but may have felt like they "lost". They may view this as a negative negotiation outcome, but our model would rate this as a successful negotiation (agreement reached).

The size and nature of the data set we used may limit the applicability of the methodology we used. The set has 7,199 individual utterances and 147,220 words, which should provide enough data for the speech act classification model which works on individual utterances. However, the second model, which predicts negotiation outcome, only has 20 negotiations as units of study. A larger corpus of negotiations would certainly lend more credibility to the findings. Also, the data consisted solely of divorce negotiations. The reported results may not generalize well to other kinds of negotiations such as nation-to-nation, business-to-business, or labor negotiations.

To date, all of our testing has occurred in post-collection analysis of the negotiation. We have not tested the process in on-going, real-time negotiations.

#### 4.2 Future Work

More study is needed on making outcome predictions at intervals in the negotiation. The negotiations could be analyzed a variety of points in time and the scores could be tracked. This staged approach would allow a broader view of the negotiation's direction and allow for vectors to be calculated for entire negotiation sections as well

as relative movement towards or away from consensus. It would also allow further study on making the stages of negotiation definable by speech act trajectory. It may be possible to further refine the scores to predict and define discrete negotiation stages, which would aid negotiators in knowing when to employ certain tactics and to reduce unexpected results. Finally, more needs to be done in the area of visualization. This technology could be linked to a negotiation decision support system that would quickly present necessary trajectory information in real-time.

## 5 Conclusion

This paper reports the results of a study attempting to classify the outcome of 20 negotiation transcripts using a novel two-step classification model. The use of this model begins with extraction of syntactical and semantic language features from individual utterances in a negotiation. These features are fed into a machine-learning algorithm, which classifies the utterance along a scale of integrative and distributive speech acts. These scores are aggregated across the negotiation with a greater weighting on the end of the negotiation and are fed into another classification algorithm, which makes the final classification of successful (agreement reached) or unsuccessful (agreement not reached) negotiation outcome. The model's resulting accuracy is good when compared to other efforts at negotiation outcome classification, and the model provides a mediating variable that gives the user reasoning behind the classification. The long-standing use of integrative and distributive labels to describe negotiation strategies lends support to the models use of the mediating variable. Future work involving larger data sets, data sets of other types of negotiations, and partial negotiations should aid in future modifications to the model and help further the understanding of language and negotiations.

**Acknowledgments** For use of the divorce negotiations transcripts, the authors wish to thank William Donohue.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Austin J L (1962) How to do things with words. Harvard University Press, Cambridge, MA
- Bartos OJ (1964) A model of negotiation and the recency effect. *Sociometry* 27(3):311–326
- Carberry S, Lambert L (1999) A process model for recognizing communicative acts and modeling negotiation subdialogues. *Comput Linguist* 25(1):1–53
- Chang MK, Woo CC (1994) A speech-act-based negotiation protocol: design, implementation, and test use. *ACM Trans Inf Syst (TOIS)* 12(4):360–382
- Donohue WA, Ramesh C, Kaufman G, Smith R (1991) Crisis bargaining: tracking relational paradox in hostage negotiation. *Int J Group Tens* 21:133–145
- Donohue WA, Roberto J (1996) An empirical examination of three models of integrative and distributive bargaining. *Int J Confl Manag* 7(3):209–229



- Freitas AA, Wieser DC, Apweiler R (2010) On the importance of comprehensible classification models for protein function prediction. *IEEE/ACM Trans Comput Biol Bioinf (TCBB)* 7:172–182
- Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL (2006) *Multivariate data analysis*. Pearson Prentice Hall, Upper Saddle River, NJ
- Hall MA (1998) Correlation-based feature subset selection for machine learning. Hamilton, New Zealand
- Hall MA, Smith LA (1998) Practical feature subset selection for machine learning. Paper presented at the proceedings of the 21st Australasian computer science conference, Perth
- Kersten GE, Noronha SJ (1999) WWW-based negotiation support: design, implementation, and use. *Decis Support Syst* 25(2):135–154
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- Leech G, Svartvik J (2002) *A communicative grammar of English*. Pearson Longman, Upper Saddle River, NJ
- Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ (2007) *The development and psychometric properties of LIWC 2007*. Austin, TX: LIWC.net
- Pennebaker JW, Mayne TJ, Francis ME (1997) Linguistic predictors of adaptive bereavement. *J Pers Soc Psychol* 72:863–871
- Putnam LL (1990) Reframing integrative and distributive bargaining: a process perspective. *Res Negot Organ* 2:3–30
- Putnam LL, Jones TS (1990) The role of communication in bargaining. *Hum Commun Res* 8(3):262–280
- Searle JR (1979) A taxonomy of illocutionary acts. In: *Expression and meaning: studies in the theory of speech acts*. Cambridge. Cambridge University Press, UK, pp 1–29
- Sebenius JK (1992) Negotiation analysis: a characterization and review. *Manag Sci* 38(1):18–38
- Shevade SK, Keerthi SS, Bhattacharyya C, Murthy KKR (2000) Improvements to the SMO algorithm for SVM regression. *IEEE Trans Neural Netw* 11(5):1188–1193
- Sinaceur M, Neale MA (2005) Not all threats are created equal: how implicitness and timing affect the effectiveness of threats in negotiations. *Group Decis Negotiat* 14(1):63–85
- Smola AJ, Schoelkopf B. (2004) A tutorial on support vector regression. *Stat Comput* 14:199–222
- Sokolova M, Szpakowicz S (2005) Analysis and classification of strategies in electronic negotiations. In: Kegal B, Lapalme G (eds) *Advances in artificial intelligence (Proceedings of the 18th conference of the Canadian Society for Computational Studies of Intelligence, Canadian AI 2005, held in Victoria, Canada in May 2005)*. Springer, New York, NY, pp 145–157
- Sokolova M, Szpakowicz S. (2007) Strategies and language trends in learning success and failure of negotiation. *Group Decis Negotiat* 16(6):469–484
- Stolcke A, Ries K, Coccaro N, Shriberg E, Bates R, Jurafsky D et al (2000) Dialogue act modeling for automatic tagging and recognition of conversational speech. *Comput Linguist* 26(3):339–373
- Twitchell DP, Nunamaker JF (2004) Speech act profiling: a probabilistic method for analyzing persistent conversations and their participants. Paper presented at the thirty-seventh annual Hawaii international conference on system sciences (CD-ROM), Big Island, Hawaii
- Walton RE, McKersie RB (1991) *A behavioral theory of labor negotiations: an analysis of a social*. Cornell University Press, Ithaca, NY
- Whissell C (1989) The dictionary of affect in language. In: Plutchnik R, Kellerman H (eds) *Emotion: theory, research and experience*. Academic Press, New York, NY p 113
- Whissell C, Fournier M, Pelland R, Weir D, Makarec K (1986) A dictionary of affect in language: IV. Reliability, validity, and applications. *Percept Mot Skills* 62:875–888
- Winograd T, Flores F (1986) *Understanding computers and cognition: a new foundation for design*. Ablex Publishing, Norwood, NJ
- Witten IE, Frank E (2005) *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, CA
- Zhou L, Burgoon JK, Nunamaker JF, Twitchell DP (2004a) Automated linguistics based cues for detecting deception in text-based asynchronous computer-mediated communication: an empirical investigation. *Group Decis Negotiat* 13(1):81–106
- Zhou L, Twitchell DP, Qin T, Burgoon JK, Nunamaker JF (2004b) Toward the automatic prediction of deception—an empirical comparison of classification methods. *J Manag Inf Syst* 20(4):139–166